A

Please type a plus sign (+) inside this box → [+]

PTO/SB/05 (4/98)
Approved for use through 09/30/2000. OMB 0651-0032
Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

UTILITY PATENT APPLICATION TRANSMITTAL

(Only for new nonprovisional applications under 37 C.F.R. § 1.53(b))

Attorney Docket No.	Mo-5176/LeA 33,020
First Inventor or Application Identifier	ADAMCZEWSKI ET AL
Title	NUCLEIC ACIDS WHICH ENCODE INSECT....
Express Mail Label No.	EJ564485416US

APPLICATION ELEMENTS

See MPEP chapter 600 concerning utility patent application contents.

ADDRESS TO: Assistant Commissioner for Patents, Box Patent Application, Washington, DC 20231

1. [X] * Fee Transmittal Form *(e.g., PTO/SB/17)* *(Submit an original and a duplicate for fee processing)*
2. [X] Specification [Total Pages 24] *(preferred arrangement set forth below)*
   - Descriptive title of the Invention
   - Cross References to Related Applications
   - Statement Regarding Fed sponsored R & D
   - Reference to Microfiche Appendix
   - Background of the Invention
   - Brief Summary of the Invention
   - Brief Description of the Drawings *(if filed)*
   - Detailed Description
   - Claim(s)
   - Abstract of the Disclosure
3. [1] Drawing(s) *(35 U.S.C. 113)* [Total Sheets 1]
4. Oath or Declaration [Total Pages]
   - a. [ ] Newly executed (original or copy)
   - b. [ ] Copy from a prior application (37 C.F.R. § 1.63(d)) *(for continuation/divisional with Box 16 completed)*
     - i. [ ] DELETION OF INVENTOR(S) Signed statement attached deleting inventor(s) named in the prior application, see 37 C.F.R. §§ 1.63(d)(2) and 1.33(b).

** NOTE FOR ITEMS 1 & 13 IN ORDER TO BE ENTITLED TO PAY SMALL ENTITY FEES, A SMALL ENTITY STATEMENT IS REQUIRED (37 C.F.R. § 1.27), EXCEPT IF ONE FILED IN A PRIOR APPLICATION IS RELIED UPON (37 C.F.R. § 1.28).*

5. [ ] Microfiche Computer Program *(Appendix)*
6. Nucleotide and/or Amino Acid Sequence Submission *(if applicable, all necessary)*
   - a. [X] Computer Readable Copy
   - b. [X] Paper Copy (identical to computer copy)
   - c. [X] Statement verifying identity of above copies

ACCOMPANYING APPLICATION PARTS

7. [ ] Assignment Papers (cover sheet & document(s))
8. [ ] 37 C.F.R.§3.73(b) Statement *(when there is an assignee)* [ ] Power of Attorney
9. [ ] English Translation Document *(if applicable)*
10. [ ] Information Disclosure Statement (IDS)/PTO-1449 [ ] Copies of IDS Citations
11. [X] Preliminary Amendment
12. [X] Return Receipt Postcard (MPEP 503) *(Should be specifically itemized)*
13. [ ] * Small Entity Statement(s) *(PTO/SB/09-12)* [ ] Statement filed in prior application Status still proper and desired
14. [X] Certified Copy of Priority Document(s) *(if foreign priority is claimed)*
15. [ ] Other: .....

16. **If a CONTINUING APPLICATION,** *check appropriate box, and supply the requisite information below and in a preliminary amendment:*

[ ] Continuation [ ] Divisional [ ] Continuation-in-part (CIP) of prior application No: _____/_____

Prior application information: Examiner _____ *Group / Art Unit:* _____

For CONTINUATION or DIVISIONAL APPS only: The entire disclosure of the prior application, from which an oath or declaration is supplied under Box 4b, is considered a part of the disclosure of the accompanying continuation or divisional application and is hereby incorporated by reference. The incorporation can only be relied upon when a portion has been inadvertently omitted from the submitted application parts.

17. CORRESPONDENCE ADDRESS

[ ] *Customer Number or Bar Code Label* *(Insert Customer No. or Attach bar code label here)* or [X] *Correspondence address below*

Name	BAYER CORPORATION Patent Dept.				
Address	100 Bayer Road				
City	Pittsburgh	State	PA	Zip Code	15205
Country	US	Telephone	(412)777-2349	Fax	(412)777-5449

Name (Print/Type)	Lyndanne M. Whalen	Registration No. (Attorney/Agent)	29,457
Signature	Lyndanne M. Whalen	Date	4/30/99

Burden Hour Statement: This form is estimated to take 0.2 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, Patent and Trademark Office, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Assistant Commissioner for Patents, Box Patent Application, Washington, DC 20231.

PATENT APPLICATION
Mo-5176
LeA 33,020

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION OF)
)
MARTIN ADAMCZEWSKI ET AL)
)
SERIAL NUMBER: TO BE ASSIGNED)
)
FILED: HEREWITH)
)
TITLE: NUCLEIC ACIDS WHICH ENCODE)
INSECT ACETYLCHOLINE)
RECEPTOR SUBUNITS)

STATEMENT VERIFYING IDENTITY OF SEQUENCE LISTING

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

Applicants hereby state that the content of the paper form of the sequence listing which appears after the Abstract page at pages 1 to 22 of the enclosed application and the enclosed disk containing a sequence listing in computer readable form are one and the same.

The attached diskette is submitted in American Standard Code for Information Interchange (ASCII) text and is properly labeled for the convenience of the U.S. Patent & Trademark Office.

Respectfully submitted,

MARTIN ADAMCZEWSKI
NADJA OELLERS
THOMAS SCHULTE

By _____
Lyndanne M. Whalen
Attorney for Applicants
Reg. No. 29,457

Bayer Corporation
100 Bayer Road
Pittsburgh, Pennsylvania 15205-9741
(412) 777-2347
FACSIMILE PHONE NUMBER:
(412) 777-5449
s:\ksl\LMW1169

"Express Mail" mailing label number EJ564485416US
Date of Deposit April 30, 1999

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to the Assistant Commissioner of Patents and Trademarks, Washington, D.C. 20231

Dorothy P. Colangelo
(Name of person mailing paper or fee)

Dorothy P. Colangelo
Signature of person mailing paper or fee)

PATENT APPLICATION
Mo-5176
LeA 33,020

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION OF)
)
MARTIN ADAMCZEWSKI ET AL)
)
SERIAL NUMBER: TO BE ASSIGNED)
)
FILED: HEREWITH)
)
TITLE: NUCLEIC ACIDS WHICH ENCODE)
INSECT ACETYLCHOLINE)
RECEPTOR SUBUNITS)

PRELIMINARY AMENDMENT

Assistant Commissioner for Patents

Washington, D.C. 20231

Sir:

Upon granting a Serial Number to the enclosed application, please amend this application as follows:

IN THE SPECIFICATION:

At page 4, below line 8 and above line 10, please insert the following:

- - BRIEF DESCRIPTION OF THE DRAWING

Figure 1 is a graph illustrating the increase in intracellular calcium that occurs in cells which have been recombinantly modified as described in Example 2.- -

At page 16, line 1, delete the heading "**References**" and substitute therefor - - PRIOR ART- -.

IN THE CLAIMS:

Please cancel Claim 21.

Please rewrite Claims 1-20 to read as follows:

1. (Amended) A [N]nucleic acid which comprises a sequence selected from

(a) the sequences according to SEQ ID NO: 1, SEQ ID NO: 3 or SEQ ID NO: 5,

"Express Mail" mailing label number EJ564485416US
Date of Deposit April 30, 1999

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to the Assistant Commissioner of Patents and Trademarks, Washington, D.C. 20231

Dorothy P. Colangelo
(Name of person mailing paper or fee)

Dorothy P. Colangelo
Signature of person mailing paper or fee)

(b) partial sequences[,] which are at least 14 base pairs in length of [the] a sequence[s] defined under (a),

(c) sequences which hybridize with any of the sequences defined under (a) in 2 x SSC at 60°C[, preferably in 0.5 x SSC at 60°C, particularly preferably in 0.2 x SSC at 60°C],

(d) sequences which exhibit at least 70% identity with any of the sequences defined under (a)[,] between position 1295 and position 2195 from SEQ ID NO: 1, or between position 432 and position 1318 from SEQ ID NO: 3, or between position 154 and position 1123 from SEQ ID NO: 5,

(e) sequences which are complementary to the sequences defined under (a), and

(f) sequences which, [on account of the] due to degeneracy of [the] genetic code, encode the same amino acid sequences as the sequences defined under (a), (b), (c) and [to] (d).

2. (Amended) A [V]vector which comprises at least one nucleic acid [according to] of Claim 1.

3. (Amended) The [V]vector [according to] of Claim 2, characterized in that the nucleic acid is functionally linked to regulatory sequences which ensure [the] expression of the nucleic acid in prokaryotic or eukaryotic cells.

4. (Amended) A [H]host cell which contains a nucleic acid [according to] of Claim 1 [or a vector according to Claim 2 or 3].

5. (Amended) A [H]host cell [according to] of Claim 4, characterized in that it is a prokaryotic or a eukaryotic cell.

6. (Amended) A [H]host cell [according to] of Claim 5, characterized in that the prokaryotic cell is E.coli.

7. (Amended) A [H]host cell [according to] of Claim 5, characterized in that the eukaryotic cell is a mammalian cell or an insect cell.

8. (Amended) A [P]polypeptide which is encoded by a nucleic acid [according to] of Claim 1.

9. (Amended) An [A]acetylcholine receptor which comprises at least one polypeptide [according to] of Claim 8.

10. (Amended) A [P]process for preparing a polypeptide [according to Claim 8, which comprises] encoded by a nucleic acid of Claim 1 comprising

(a) culturing a host cell [according to one of Claims 4 to 7] containing a nucleic acid of Claim 1 or a vector comprising at least one nucleic acid of Claim 1 under conditions which ensure [the] expression of the nucleic acid [according to] of Claim 1, and

(b) isolating the polypeptide from the cell or the culture medium.

11. (Amended) An [A]antibody which reacts specifically with the polypeptide [according to] of Claim 8 [or the receptor according to Claim 9].

12. (Amended) A [T]transgenic invertebrate which contains a nucleic acid [according to] of Claim 1.

13. (Amended) The [T]transgenic invertebrate [according to] of Claim 12, characterized in that it is Drosophila melanogaster or Caenorhabditis elegans.

14. (Amended) A [P]process for producing a transgenic invertebrate [according to Claim 12 or 13,] which comprises introducing a nucleic acid [according to] of Claim 1 or a vector comprising at least one nucleic acid of Claim 1 [according to Claim 2 or Claim 3].

15. (Amended) The [T]transgenic progeny of an invertebrate [according to] Claim 12 [or 13].

16. (Amended) A [P]process for preparing a nucleic acid [according to] of Claim 1[, which comprises the following steps:] comprising

(a) carrying out an entirely chemical synthesis [in a manner known per se,] or

(b) chemically synthesizing an oligonucleotide[s], labelling the oligonucleotide[s], hybridizing the oligonucleotide[s] to the DNA of an insect cDNA library, selecting a positive clone[s] and isolating the hybridizing DNA from a positive clone[s], or

(c) chemically synthesizing an oligonucleotide[s] and amplifying the target DNA by means of PCR.

17. (Amended) The [R]regulatory region which naturally controls transcription of a nucleic acid [according to] of Claim 1 in insect cells and ensures specific expression.

18. (Amended) A [P]process for discovering novel active compounds for plant protection, in particular, compounds which alter the conducting properties of an acetylcholine receptor[s according to Claim 9] made up of at least one polypeptide encoded by a nucleic acid of Claim 1, which comprises the following steps:

(a) providing a host cell [according to one of Claims 4 to 7] containing a nucleic acid of Claim 1 or a vector comprising at least one nucleic acid of Claim 1,

(b) culturing the host cell in the presence of at least one [a] compound [or a sample which comprises a multiplicity of compounds], and

(c) detecting altered receptor properties.

19. (Amended) A [P]process for discovering a compound which binds to an acetylcholine receptor[s according to Claim 9, which encompasses the following steps:] comprising

(a) bringing a host cell [according to one of Claims 4 to 7] containing a nucleic acid of Claim 1 or a vector comprising at least one nucleic acid of Claim 1, a polypeptide [according to Claim 8] encoded by a nucleic acid of Claim 1 or [a] an acetylcholine receptor [according to Claim 9] comprising at least one polypeptide encoded by a nucleic acid of Claim 1 into contact with [a] at least one compound [or a mixture of compounds] under conditions which permit interaction of the compound [compound(s)] with the host cell, the polypeptide or the receptor, and

(b) determining the compound [compound(s)] which bind [bind(s)] specifically to the receptor[s].

20. (Amended) A [P]process for discovering compounds which alter the expression of an acetylcholine receptor comprising at least one polypeptide encoded by a nucleic acid of Claim 1 [receptors according to Claim 9,] which comprises the following steps:

(a) bringing a host cell [according to one of Claims 4 to 7] containing a nucleic acid of Claim 1 or a vector comprising at least one nucleic acid of Claim 1 or a transgenic invertebrate containing a nucleic acid of

Claim 1 [according to Claim 11 or Claim 12] into contact with [a] at least one compound [or a mixture of compounds],

(b) determining the receptor concentration, and

(c) determining the [compound(s)] compound which specifically [influence(s)] influences the expression of the receptor.

Please add the following new Claims 22-34:

--22. The nucleic acid of Claim 1 which comprises a sequence that hybridizes with a sequence defined under (a) in 0.5 x SSC at 60°C.

23. The nucleic acid of Claim 1 which comprises the sequence that hybridized with a sequence defined in (a) in 0.2 x SSC at 60°C.

24. A host cell containing the vector of Claim 2.

25. A host cell containing the vector of Claim 3.

26. The host cell of Claim 24 that is a prokaryotic or a eukaryotic cell.

27. The host cell of Claim 25 that is prokaryotic or a eukaryotic cell.

28. The host cell of Claim 26 that is an E. coli cell.

29. The host cell of Claim 27 that is an E. coli cell.

30. The host cell of Claim 26 that is a mammalian or an insect cell.

31. The host cell of Claim 27 that is a mammalian or an insect cell.

32. An antibody which reacts specifically with the acetylcholine receptor of Claim 9.

33. A transgenic progeny of the invertebrate of Claim 13.--

REMARKS

The specification has been amended at page 4 to insert a Brief Description of the Drawing. This description corresponds to that given at page 8, lines 19-20 of the specification.

The specification has also been amended at page 16 to change the heading from "References" to "Prior Art" to more accurately reflect the status of these disclosures.

Claim 21 has been cancelled.

Claims 1-20 have been rewritten to place them in better grammatical form and to remove the multiple dependencies which occurred therein.

New Claims 22 and 23 are directed to subject matter deleted from original Claim 1.

New Claims 24 and 25 are directed to subject matter deleted from original Claim 4.

New Claims 26 and 27 are directed to subject matter deleted from original Claim 5.

New Claims 28 and 29 are directed to subject matter deleted from original Claim 6.

New Claims 30 and 31 are directed to subject matter deleted from original Claim 7.

New Claim 32 is directed to subject matter deleted from original Claim 11.

New Claim 33 is directed to subject matter deleted from original Claim 14.

An Action on the merits of this case is respectfully requested.

Respectfully submitted,

MARTIN ADAMCZEWSKI
NADJA OELLERS
THOMAS SCHULTE

By _____
Lyndanne M. Whalen
Attorney for Applicants
Reg. No. 29,457

Bayer Corporation
100 Bayer Road
Pittsburgh, Pennsylvania 15205-9741
(412) 777-2347
FACSIMILE PHONE NUMBER:
(412) 777-5449
s:\ksl\LMW1168

Nucleic acids which encode insect acetylcholine receptor subunits

The invention relates, in particular, to nucleic acids which encode insect acetylcholine receptor subunits.

Nicotinic acetylcholine receptors are ligand-regulated ion channels which are of importance in neurotransmission in the animal kingdom. The binding of acetylcholine or other agonists to the receptor induces a transient opening of the channel and allows cations to flow through. It is assumed that a receptor consists of five subunits which are grouped around a pore. Each of these subunits is a protein which consists of an extracellular N-terminal moiety followed by three transmembrane regions, an intracellular moiety, a fourth transmembrane region and a short extracellular C-terminal moiety (Changeux et al. 1992).

Acetylcholine receptors are especially well investigated in vertebrates. In this context, three groups can be distinguished on the basis of their anatomical location and their functional properties (conducting properties of the channel, desensitization, and sensitivity towards agonists and antagonists and also towards toxins such as α-bungarotoxin). The classification correlates with the molecular composition of the receptors. There are heterooligomeric receptors having the subunit composition $\alpha_2\beta\gamma\delta$, which are found in muscle (Noda et al. 1982, Claudio et al. 1983, Devillers-Thiery et al. 1983, Noda et al. 1983a, b), heterooligomeric receptors which contain subunits from the α2 - α6 and ß2 - ß4 groups and which are found in the nervous system (Wada et al. 1988, Schoepfer et al. 1990, Cockcroft et al. 1991, Heinemann et al. 1997), and also homooligomeric receptors which contain subunits from the α7 - α9 group and which are likewise found in the nervous system (Lindstrom et al. 1997, Elgoyhen et al. 1997). This classification is also supported by an examination of the relatedness of the gene sequences of the different subunits. Typically, the sequences of functionally homologous subunits from different species are more similar to each other than are sequences of subunits which are from different groups but from the

same species. Thus, the rat muscle α subunit, for example, exhibits 78% amino acid identity and 84% amino acid similarity with that of the electric ray Torpedo californica but only 48% identity and 59% similarity with the rat α2 subunit (hetero-oligomeric, neuronal) and 36% identity and 45% similarity with the rat α7 subunit (homooligomeric, neuronal). Furthermore, the gene sequences of all the known acetylcholine receptor subunits are to a certain extent similar not only to each other but also to those of some other ligand-regulated ion channels (e.g. the serotonin receptors of the $5HT_3$ type, the GABA-regulated chloride channels and the glycine-regulated chloride channels). It is therefore assumed that all these receptors are descended from one common precursor and they are classified into one supergene family (Ortells et al. 1995).

In insects, acetylcholine is the most important excitatory neurotransmitter of the central nervous system. Accordingly, acetylcholine receptors can be detected electrophysiologically in preparations of insect central nervous system ganglia. The receptors are detected both in postsynaptic and presynaptic nerve endings and in the cell bodies of interneurones, motor neurones and modulatory neurones (Breer et al. 1987, Buckingham et al. 1997). Some of the receptors are inhibited by α-bungarotoxin while others are insensitive (Schloß et al. 1988). In addition, the acetylcholine receptors are the molecular point of attack for important natural (e.g. nicotine) and synthetic insecticides (e.g. chloronicotinyls).

The gene sequences of a number of insect nicotinic acetylcholine receptors are already known. Thus, the sequences of five different subunits have been described in Drosophila melanogaster (Bossy et al. 1988, Hermanns-Borgmeyer et al. 1986, Sawruk et al. 1990a, 1990b, Schulz et al. Unpublished, EMBL accession number Y15593), while five have likewise been described in Locusta migratoria (Stetzer et al. unpublished, EMBL accession numbers AJ000390 - AJ000393), one has been described in Schistocerca gregaria (Marshall et al. 1990), two have been described in Myzus persicae (Sgard et al. unpublished, EMBL accession number X81887 and X81888), and one has been described in Manduca sexta (Eastham et al. 1997). Fur-

thermore, a number of partial gene sequences from Drosophila melanogaster have been characterized as so-called expressed sequence tags (Genbank accession numbers AA540687, AA698155, AA697710, AA697326). The fact that individual sequences are very similar to those from other insects suggests that these subunits are functional homologues.

It is of great practical importance to make available new insect acetylcholine receptor subunits, for example for the purpose of searching for novel insecticides, with those subunits which differ from the known subunits to a greater extent than is the case between functional homologues being particularly of interest.

The present invention is consequently based, in particular, on the object of making available nucleic acids which encode novel insect acetylcholine receptor subunits.

This object is achieved by the provision of nucleic acids which comprise a sequence selected from

(a) the sequences according to SEQ ID NO: 1, SEQ ID NO: 3 or SEQ ID NO: 5,

(b) part sequences of the sequences defined in (a) which are least 14 base pairs in length,

(c) sequences which hybridize to the sequences defined in (a) in 2 x SSC at 60°C, preferably in 0.5 x SSC at 60°C, particularly preferably in 0.2 x SSC at 60°C (Sambrook et al. 1989),

(d) sequences which exhibit at least 70% identity with the sequences defined in (a), between position 1295 and position 2195 in the case of SEQ ID NO: 1, or between position 432 and position 1318 in the case of SEQ ID NO: 3, or between position 154 and position 1123 in the case of SEQ ID NO: 5,

(e) sequences which are complementary to the sequences defined in (a), and

(f) sequences which, because of the degeneracy of the genetic code, encode the same amino acid sequences as the sequences defined in (a) to (d).

The degree of identity of the nucleic acid sequences is preferably determined using the GAP program from the GCG program package, Version 9.1 with standard settings (Devereux et al. 1984).

The present invention is based on the surprising finding that insects possess genes which encode subunits of, in particular, homooligomeric acetylcholine receptors.

The invention furthermore relates to vectors which contain at least one of the novel nucleic acids. All the plasmids, phasmids, cosmids, YACs or artificial chromosomes which are used in molecular biological laboratories can be used as vectors. These vectors can be linked to the usual regulatory sequences for the purpose of expressing the novel nucleic acids. The choice of such regulatory sequences depends on whether prokaryotic or eukaryotic cells, or cell-free systems, are used for the expression. The SV40, adenovirus or cytomegalovirus early or late promoter, the lac system, the trp system, the main operator and promoter regions of phage lambda, the control regions of the fd coat protein, the 3-phosphoglycerate kinase promoter, the acid phosphatase promoter and the yeast α-mating factor promoter are examples of expression control sequences which are particularly preferred.

In order to be expressed, the nucleic acids according to the invention can be introduced into suitable host cells. Both prokaryotic cells, preferably E.coli, and eukaryotic cells, preferably mammalian or insect cells, are suitable for use as host cells. Other examples of suitable unicellular host cells are: Pseudomonas, Bacillus, Streptomyces, yeasts, HEK-293, Schneider S2, CHO, COS1 and COS7 cells, plant cells in cell culture and also amphibian cells, in particular oocytes.

The present invention also relates to polypeptides which are encoded by the nucleic acids according to the invention and also the acetylcholine receptors, preferably homooligomeric acetylcholine receptors, which are synthesized from them.

In order to prepare the polypeptides which are encoded by the nucleic acids according to the invention, host cells which contain at least one of the nucleic acids according to the invention can be cultured under suitable conditions. After that, the desired polypeptides can be isolated from the cells or the culture medium in a customary manner.

The invention furthermore relates to antibodies which bind specifically to the above-mentioned polypeptides or receptors. These antibodies are prepared in the customary manner. For example, such antibodies can be produced by injecting a substantially immunocompetent host with a quantity of an acetylcholine receptor polypeptide, or a fragment thereof, according to the invention which is effective for producing antibodies, and subsequently isolating these antibodies. Furthermore, an immortalized cell line which produces monoclonal antibodies can be obtained in a manner known per se. Where appropriate, the antibodies can be labelled with a detection reagent. Preferred examples of such a detection reagent are enzymes, radioactively labelled elements, fluorescent chemicals or biotin. Instead of the complete antibody, use can also be made of fragments which possess the desired specific binding properties.

The nucleic acids according to the invention can be used, in particular, for producing transgenic invertebrates. These latter can be employed in test systems which are based on an expression of the receptors according to the invention, or variants thereof, which differs from that of the wild type. In addition, this includes all transgenic invertebrates in which a change in the expression of the receptors according to the invention, or their variants, occurs as the result of modifying other genes or gene control sequences (promoters).

The transgenic invertebrates are produced, for example, in Drosophila melanogaster by means of P element-mediated gene transfer (Hay et al., 1997) or in Caenorhabditis elegans by means of transposon-mediated gene transfer (e.g. using Tc1, Plasterk, 1996).

The invention also consequently relates to transgenic invertebrates which contain at least one of the nucleic acid sequences according to the invention, preferably to transgenic invertebrates of the species Drosophila melanogaster or Caenorhabditis elegans, and to their transgenic progeny. Preferably, the transgenic invertebrates contain the receptors according to the invention in a form which differs from that of the wild type.

The nucleic acids according to the invention can be prepared in the customary manner. For example, the nucleic acid molecules can be synthesized entirely chemically. In addition, only short segments of the sequences according to the invention can be synthesized chemically and these oligonucleotides can be labelled radioactively or with a fluorescent dye. The labelled oligonucleotides can be used to screen cDNA libraries prepared from insect mRNA. Clones which hybridize to the labelled oligonucleotides ("positive clones") are selected for isolating the relevant DNA. After the isolated DNA has been characterized, the nucleic acids according to the invention are readily obtained.

The nucleic acids according to the invention can also be prepared by means of PCR methods using chemically synthesized oligonucleotides.

The nucleic acids according to the invention can be used for isolating and characterizing the regulatory regions which occur naturally adjacent to the coding region. Consequently, the present invention also relates to these regulatory regions.

The nucleic acids according to the invention can be used to identify novel active compounds for plant protection, such as compounds which, as modulators, in particular as agonists or antagonists, alter the conducting properties of the acetylcholine receptors according to the invention. For this, a recombinant DNA molecule, which encompasses

at least one nucleic acid according to the invention, is introduced into a suitable host cell. The host cell is cultured, in the presence of a compound or a sample which comprises a multiplicity of compounds, under conditions which permit expression of the receptors according to the invention. A change in the receptor properties can be detected, as described below in Example 2. Using this approach, it is possible to discover insecticidal substances.

The nucleic acids according to the invention also make it possible to discover compounds which bind to the receptors according to the invention. These compounds can likewise be used as insecticides on plants. For example, host cells which contain the nucleic acid sequences according to the invention and express the corresponding receptors or polypeptides, or the gene products themselves, are brought into contact with a compound or a mixture of compounds under conditions which permit the interaction of at least one compound with the host cells, receptors or the individual polypeptides.

Host cells or transgenic invertebrates with contain the nucleic acids according to the invention can also be used to discover substances which alter the expression of the receptors.

The above-described nucleic acids, vectors and regulatory regions according to the invention can additionally be used for discovering genes which encode polypeptides which are involved in the synthesis, in insects, of functionally similar acetylcholine receptors. According to the present invention, functionally similar receptors are understood as being receptors which encompass polypeptides which, while differing in their amino acid sequences from the polypeptides described in this present publication, essentially possess the same functions.

Comments on the sequence listing and the figures:

SEQ ID NO: 1 shows the nucleotide sequence of the isolated Da7 cDNA, beginning with position 1 and ending with position 2886. SEQ ID NO: 1 and SEQ ID NO: 2 also show the amino acid sequences of the protein deduced from the Da7 cDNA sequence.

SEQ ID NO: 3 shows the nucleotide sequence of the isolated Hva7-1 cDNA, beginning with position 1 and ending with position 3700. SEQ ID NO: 3 and SEQ ID NO: 4 also show the amino acid sequences of the protein deduced from the Hva7-1 cDNA sequence.

SEQ ID NO: 5 shows the nucleotide sequence of the isolated Hva7-2 cDNA, beginning with position 1 and ending with position 3109. SEQ ID NO: 5 and SEQ ID NO: 6 also show the amino acid sequences of the protein deduced from the Hva7-2 cDNA sequence.

Figure 1 shows the increase in intracellular calcium which occurs in cells which have been recombinantly modified as described in Example 2 following the addition of nicotine. Cells were loaded with Fura-2-acetoxymethyl ester (5 - 10 µM in serum-free minimal essential medium containing 1% bovine serum albumin and 5 mM calcium chloride), washed with Tyrode solution buffered with N-(2-hydroxyethyl)piperazine-N'-(2-ethanesulphonic acid) (5 mM HEPES) and alternately illuminated, under a fluorescence microscope (Nikon Diaphot) with light of 340 nm and 380 nm wavelength. A measurement point corresponds to a pair of video images at the two wavelengths (exposure time per image, 100 ms). The time interval between two measurement points is 3 s. After 8 images had been taken (measurement point 4.0), nicotine was added to a final concentration of 500 µM and the measurement series was continued. The fluorescence intensity of the cells when illuminated

with light of 380 nm wavelength was divided by the corresponding intensity at 340 nm, thereby giving the ratio.

Examples:

Example 1

Isolating the described polynucleotide sequences

Polynucleotides were manipulated using standard methods of recombinant DNA technology (Sambrook, et al., 1989). The bioinformatic processing of nucleotide and protein sequences was carried out using the GCG program package Version 9.1 (GCG Genetics Computer Group, Inc., Madison Wisconsin, USA).

Partial polynucleotide sequences

Sequence comparisons ("Clustalw") were used to identify regions, from which degenerate oligonucleotides were deduced by backtranslating the codons, of protein sequences from genes whose ability to form homooligomeric acetylcholine receptors was known. In all, 5 such oligonucleotide pairs were selected for the polymerase chain reaction (PCR). Only one combination (see below) gave a product both from Heliothis cDNA and from Drosophila cDNA.

RNA was isolated from whole Heliothis virescens embryos (shortly before hatching) using Trizol reagent (Gibco BRL, in accordance with the manufacturer's instructions). The same procedure was adopted with Drosophila embryos (24 h at 25°C). 10 µg of these RNAs were employed in a first cDNA strand synthesis (Superscript Preamplification System for first cDNA strand synthesis, Gibco BRL, in accordance with the manufacturer's instructions, reaction temperature 45°C).

Subsequently, 1/100 of the abovementioned first-strand cDNA was in each case employed in a polymerase chain reaction (PCR) using the oligonucleotides alpha7-1s: (5'-GAYGTIGAYGARAARAAYCA-3') and alpha7-2a: (5'-CYYTCRTCIGCRCTRTTRTA-3') (recombinant Taq DNA polymerase, Gibco BRL). The PCR parameters were as follows: Hva7-1 and Hva7-2: 94°C, 2 min; 35 times (94°C, 45 s; 50°C, 30 s; 72°C, 60 s) and also Da7: 96°C, 2 min; 35 times (96°C, 45 s; 50°C, 30 s; 72°C, 60 s). In each case, this resulted in a dectable band of approx. 0.2 kb in an agarose gel (1%), both in the case of Drosophila cDNA and in the case of Heliothis cDNA. After the DNA fragments had been subcloned by means of SrfScript (Stratagene), and their sequences had been determined, it turned out that two different DNA fragments had been amplified from Heliothis cDNA; these were 228-11 = Hva7-1 (partial, containing 165 bp) and 228-8 = Hva7-2 (partial, containing 171 bp). Only one DNA fragment was isolated from Drosophila cDNA; this was 248-5 = Da7 (partial, containing 150 bp).

Isolating poly A-containing RNA from Heliothis virescens tissue and constructing the cDNA libraries

The RNA for cDNA library I was isolated from whole Heliothis virescens embryos (shortly before hatching) using Trizol reagent (Gibco BRL, in accordance with the manufacturer's instructions). The RNA for cDNA library II was isolated from whole head ganglia from 500 Heliothis virescens larvae (stages 4-5) usings Trizol reagent (Gibco BRL, in accordance with the manufacturer's instructions). The poly A-containing RNAs were then isolated from these RNAs by purifying with Dyna Beads 280 (Dynal). 5 µg of these poly A-containing RNAs were subsequently employed in constructing cDNA libraries I and II using the λ-ZAPExpress vector (cDNA Synthesis Kit, ZAP-cDNA Synthesis Kit and ZAP-cDNA Gigapack III Gold Cloning Kit, all from Stratagene). In a departure from the manufacturer's instructions, Superscript Reverse Transcriptase (Gibco BRL) was used for synthesizing the cDNA at a synthesis temperature of 45°C. In addition, radioactively labelled deoxynucleoside triphosphates were not added. Furthermore, the synthetisized cDNAs were not frac-

tionated through the gel filtration medium contained in the kit but instead through Size Sep 400 Spun Columns (Pharmacia).

Complete polynucleotide sequences

Apart from the first screening round when isolating the Hva7-1 clone, all the screens were carried out using the DIG system (all reagents and consumables from Boehringer Mannheim, in accordance with the instructions in "The DIG System User's Guide for Filter Hybridization", Boehringer Mannheim). The DNA probes employed were prepared by means of PCR using digoxigenin-labelled dUTP. The hybridizations were carried out at 42°C overnight in DIG Easy Hyb (Boehringer Mannheim). Labelled DNA was detected on nylon membranes by means of chemiluminescence (CDP-Star, Boehringer Mannheim) using X-ray films (Hyperfilm MP, Amersham). Initial partial sequencing of the isolated gene library plasmids was carried out, for identification purposes, using T3 and T7 primers (ABI Prism Dye Terminator Cycle Sequencing Kit, ABI, using an ABI Prism 310 Genetic Analyzer). The complete polynucleotide sequences in Hva7-1, Hva7-2 and Da7 were determined, as a commissioned sequencing carried out by Qiagen, Hilden, by means of primer walking using cycle sequencing.

a. Isolating the Da7 clone

10^6 phages from a Drosophila melanogaster cDNA library in λ phages (Canton-S embryo, 2-14 hours, in Uni-ZAP XR vector, Stratagene) were screened using DIG-labelled 248-5 as the probe (in accordance with the manufacturer's (Stratagene) instructions). The maximum stringency when washing the filters was: 0.2 x SSC; 0.1% SDS; 42°C; 2 x 15 min. One clone (clone 432-1) was isolated whose insert had a size of 2940 bp (Da7, SEQ ID NO: 1). The largest open reading frame of this sequence begins at position 372 of the depicted sequence and ends at position 1822. The 770 amino acids polypeptide which is deduced from this (SEQ ID NO: 2) has a calculated molecular weight of 87.01 kD.

b. Isolating the Hva7-1 clone

10^6 phages from the Heliothis virescens embryo cDNA library (library I) were included in the screening. The first of three screening rounds took place using α-^{32}P-labelled 228-11 DNA as the probe. The probe was hybridized to the filters in Quickhyb (Stratagene) at 68°C for one hour. The filters were then washed twice, for 15 min on each occasion, at room temperature in 2 x SSC; 0.1% SDS and twice, for 30 min on each occassion, at 42°C in 0.1xSSC; 0,1% SDS. Hybridized probes were detected by means of autoradiography, at -80°C overnight, using XR X-ray films (Kodak) and employing intensifying screens (Amersham). The two further screening rounds were carried out using the DIG System (Boehringer Mannheim).

The clone 241-5, which was isolated in this screen, contained an insert of 3630 bp. This insert (Hva7-1, SEQ ID NO: 3) possesses a longest open reading frame which begins at position 335 of the depicted nucleic acid sequence and ends at position 1821. The 496 amino acids polypeptide which is deduced from this (SEQ ID NO: 4) has a calculated molecular weight of 56.36 kD.

c. Isolating the Hva7-2 clone

10^6 phages from the Heliothis virescens ganglia cDNA library (library II) were included in the screening. Dig-labelled 228-8 DNA was used as the probe. The maximum stringency when washing the filters was: 0.1 x SSC; 0.1% SDS; 42°C; 2 x 15 min.

The clone 241-5, which was isolated in this screen, contained an insert of 3630 bp. This insert (Hva7-2, SEQ ID NO: 5) possesses a longest open reading frame which begins at position 95 of the depicted nucleic acid sequence and ends at position 1598. The 501 amino acids polypeptide which is deduced from this (SEQ ID NO: 6) has a calculated molecular weight of 56.71 kD.

Example 2

Generating the expression constructs

a. Da7

The sequence region from position 372 to position 2681 of SEQ ID NO: 1 was amplified by means of a polymerase chain reaction (PCR). Deoxyoligonucleotides having the sequences GCGAATTCACCACCATGAAAAATGCACAACTG and CGAGACAATAATATGTGGTGCCTCGAG were used for this. The Pfu polymerase from Stratagene was used as the DNA polymerase in accordance with the manufacturer's instructions. Following the amplification, the segment which had been generated was digested with the restriction endonucleases Eco RI and Xho I and cloned into a vector, i.e. pcDNA3.1/Zeo (Invitrogen), which had likewise been digested with Eco RI and Xho I.

b. Hva7-1

The sequence region from position 335 to position 1822 from SEQ ID NO: 3 was amplified by means of a polymerase chain reaction (PCR). Deoxyoligonucleotides having the sequences GCAAGCTTACCACCATGGGAGGTAGAGCTAGACGCTCGCAC and GCCTCGAGCGACACCATGATGTGTGGCGC were used for this. The Pfu polymerase from Stratagene was used as the DNA polymerase in accordance with the manufacturer's instructions. Following amplification, the generated segment was digested with the restriction endonucleases HindIII and Xho I and cloned into a vector, i.e. pcDNA3.1/Zeo (Invitrogen), which had likewise been digested with HindIII and Xho I.

c. Hva7-2

The sequence region from position 95 to position 1597 from SEQ ID NO: 5 was amplified by means of a polymerase chain reaction (PCR). Deoxyoligonucleotides having the sequences GCAAGCGCCGCTATGGCCCCTATGTTG and TTGCACGATGATATGCGGTGCCTCGAGCG were used for this. The Pfu polymerase from Stratagene was used as the DNA polymerase in accordance with the manufacturer's instructions. Following amplification, the generated segment was digested with the restriction endonucleases HindIII and Xho I and cloned into a vector, i.e. pcDNA3.1/Zeo (Invitrogen), which had likewise been digested with HindIII and Xho I.

d.Hva7-1 / $5HT_3$ and Hva7-2 / $5HT_3$ chimaeras

The region from position 335 to position 1036 from SEQ ID NO: 3 (Hva7-1/$5HT_3$ chimaera) and the region from position 95 to position 763 from SEQ ID NO: 5 (Hva7-2/$5HT_3$ chimaera) was in each case fused to the region from position 778 to position 1521 from the Mus musculus 5-HT_3 receptor cDNA (sequence in EMBL database: M774425) using the method of overlap extension (Jespersen et al. 1997). The two fragments were subsequently cloned into the pcDNA3.1/Zeo vector by means of TA cloning (Invitrogen, in accordance with the manufacturer's instructions). Constructs containing the correct orientation of the two fragments in the vector were identified by sequencing using the T7 primer (Invitrogen).

Cell culture and gene transfer

HEK293 cells, which express the α subunit of an L-type Ca channel (Zong et al. 1995, Stetzer et al. 1996), were cultured in Dulbecco's modified Eagle's medium and 10% foetal calf serum at 5% CO_2 and from 20°C to 37°C. FuGENE 6 (Boehringer Mannheim GmbH, Mannheim, Germany) was used for the gene transfer in accordance with the manufacturer's instructions. At from 24 h to 48 h after the gene trans-

fer, the cells were sown at various densities in microtitre plates. Recombinantly altered cells were selected by growth in Dulbecco's modified Eagle's medium and 10% foetal calf serum and 150 - 500 µg/ml of Zeocin/ml over a period of from 3 to 4 weeks. Individual resistant clones were analyzed as described below.

Fura-2 measurements

The alterations in the intracellular calcium concentration were measured using Fura-2. A stock solution containing 2 mM Fura-2-acetoxy methyl ester (Sigma) in dimethyl sulphoxide (DMSO) was diluted to a final concentration of 5 - 10 µM in serum-free minimal essential medium (MEM, Gibco) containing 1% bovine serum albumin and 5 mM calcium chloride. The cells were incubated for from 45 to 60 min in this solution in a microtitre plate. The cells were then washed twice in Tyrode solution buffered with N-(2-hydroxyethyl)piperazine-N'-(2-ethanesulphonic acid) (5 mM HEPES) (HEPES-buffered salt solution containing 130 mM NaCl, 5 mM KCl, 2 mM $CaCl_2$, 1 mM $MgCl_2$, 5 mM $NaHCO_3$, 10 mM glucose). 100 µl Tyrode buffer were added to the wells of the microtitre plate and the cells were illuminated alternately, under a fluorescence microscope (Nikon Diaphot), with light of 340 nm and 380 nm wavelength. A series of video images (exposure time per image 100 ms) were taken with pauses of 3 seconds and stored, as digitalized images, in an image analysis computer (Leica, Quantimet 570). After 8 images had been taken (measurement point 4.0 in Fig. 1), nicotine was added to a final concentration of 500 µM and the measurement series was continued. The fluorescence intensity of the cells when illuminating with light of 380 nm wavelength was divided by the corresponding intensity at 340 nm and in this way a ratio was formed which represents the relative increase in calcium concentration (Grynkiewicz et al. 1985).

References:

Bossy et al. (1988) Conservation of neural nicotinic acetylcholine receptors from Drosophila to vertebrate central nervous systems, EMBO J. 7, 611-618

Breer et al. (1987) Molecular properties and functions of insect acetylcholine receptors, J. Insect Physiol. 33, 771-790

Buckingham et al. (1997) Imidacloprid actions on insect neuronal acetylcholine receptors, J.Exp. Biol. 200, 2685-2692

Changeux et al. (1992) The functional architecture of the nicotinic acetylcholine receptor explored by affinity labelling and site-directed mutagenesis, Quarterly Review of Biophysics 25, 395-432

Claudio et al. (1983) Nucleotide and deduced amino acid sequences of Torpedo californica acetylcholine receptor g subunit, Proc. Natl. Acad. Sci. USA 80, 1111-1115

Devereux et al. (1984), Nucleic Acids Research 12, 387

Devillers-Thiery et al. (1983) Complete mRNA coding sequence of the acetylcholine binding α-subunit of Torpedo marmorata acetylcholine receptor: a model for the transmembrane organization of the polypeptide chain, Proc. Natl. Acad. Sci. USA 80, 2067-2071

Elgoyhen et al. (1997) US Pat. No. 5,683,912

Eastham et al. (1998) Characterisation of a nicotinic acetylcholine receptor from the insect Manduca sexta, Eur. J. Neurosci 10, 879-889

Grynkiewicz et al. (1985) A new generation of Ca2+ indicators with greatly improved fluorescence properties, J. Biol. Chem. 260, 3440-3450

Hay et al. (1997), P element insertion-dependent gene activation in the Drosophila eye, Proceedings of The National Academy of Sciences of The United States of America 94 (10), 5195-5200

Hermans-Borgmeyer et al. (1986) Primary structure of a developmentally regulated nicotinic acetylcholine receptor protein from Drosophila EMBO J. 5, 1503-1508

Heinemann et al. (1997) US Pat. No 5,591,590

Jespersen et al. (1997) Efficient Non-PCR-Mediated Overlap Extension of PCR Fragments by Exonuclease "End Polishing", Biotechniques, 23, 48-52

Lindstrom et al. (1997) US Pat. No. 5,599,709

Marshall et al. (1990) Sequence and functional expression of a single a subunit of an insect nicotinic acetylcholine receptor, EMBO J. 9, 4391-4398

Noda et al. (1982), Primary structure of α-subunit precursor of Torpedo californica acetylcholine receptor deduced from cDNA sequence, Nature 299, 793-797

Noda et al. (1983a), Primary structures of β- and δ-subunit precursor of Torpedo californica acetylcholine receptor deduced from cDNA sequences, Nature 301, 251-255

Noda et al. (1983b), Structural homology of Torpedo californica acetylcholine receptor subunits, Nature 302, 528-532

Ortells et al. (1995), Evolutionary history of the ligand-gated ion-channel superfamily of receptors, Trends in Neurosience 18, 121-127

Plasterk (1996), The Tc1/mariner transposon family, Transposable Elements, Current Topics in Microbiology and Immunology 204, 125-143

Sambrook et al. (1989), Molecular Cloning, A Laboratory Manual, 2nd ed. Cold Spring Harbour Press

Sawruk et al. (1990a), EMBO J. 9, 2671-2677 Heterogeneity of Drosophila nicotinic acetylcholine receptors: SAD, a novel developmentally regulated α-subunit

Sawruk et al. (1990b), SBD, a novel structural subunit of the Drosophila nicotinic acetylcholine receptor, shares its genomic localization with two a-subunits, FEBS Lett. 273, 177-181

Schloß et al. (1988), Neuronal acetylcholine receptors of Drosophila: the ARD protein is a component of a high-affinity α-bungarotoxin binding complex, EMBO J. 7, 2889-2984

Stetzer et al. (1996) Stable expression in HEK-293 cells of the rat α3/β4 subtype of neuronal nicotinic acetylcholine receptor, FEBS Lett. 397, 39-44

Zong et al. (1995) On the regulation of the expressed L-type calcium channel by cAMP-dependent phosphorylation, Pflügers Arch. - Eur. J. Physiol. 430, 340-347.

Patent Claims

1. Nucleic acid which comprises a sequence selected from

(a) the sequences according to SEQ ID NO: 1, SEQ ID NO: 3 or SEQ ID NO: 5,

(b) part sequences, which are least 14 base pairs in length, of the sequences defined under (a),

(c) sequences which hybridize with the sequences defined under (a) in 2 x SSC at 60°C, preferably in 0.5 x SSC at 60°C, particularly preferably in 0.2 x SSC at 60°C,

(d) sequences which exhibit at least 70% identity with the sequences defined under (a), between position 1295 and position 2195 from SEQ ID NO: 1, or between position 432 and position 1318 from SEQ ID NO: 3, or between position 154 and position 1123 from SEQ ID NO: 5,

(e) sequences which are complementary to the sequences defined under (a), and

(f) sequences which, on account of the degeneracy of the genetic code, encode the same amino acid sequences as the sequences defined under (a) to (d).

2. Vector which comprises at least one nucleic acid according to Claim 1.

3. Vector according to Claim 2, characterized in that the nucleic acid is functionally linked to regulatory sequences which ensure the expression of the nucleic acid in prokaryotic or eukaryotic cells.

4. Host cell which contains a nucleic acid according to Claim 1 or a vector according to Claim 2 or 3.

5. Host cell according to Claim 4, characterized in that it is a prokaryotic or eukaryotic cell.

6. Host cell according to Claim 5, characterized in that the prokaryotic cell is E.coli.

7. Host cell according to Claim 5, characterized in that the eukaryotic cell is a mammalian cell or an insect cell.

8. Polypeptide which is encoded by a nucleic acid according to Claim 1.

9. Acetylcholine receptor which comprises at least one polypeptide according to Claim 8.

10. Process for preparing a polypeptide according to Claim 8, which comprises

(a) culturing a host cell according to one of Claims 4 to 7 under conditions which ensure the expression of the nucleic acid according to Claim 1, and

(b) isolating the polypeptide from the cell or the culture medium.

11. Antibody which reacts specifically with the polypeptide according to Claim 8 or the receptor according to Claim 9.

12. Transgenic invertebrate which contains a nucleic acid according to Claim 1.

13. Transgenic invertebrate according to Claim 12, characterized in that it is Drosophila melanogaster or Caenorhabditis elegans.

14. Process for producing a transgenic invertebrate according to Claim 12 or 13, which comprises introducing a nucleic acid according to Claim 1 or a vector according to Claim 2 or 3.

15. Transgenic progeny of an invertebrate according to Claim 12 or 13.

16. Process for preparing a nucleic acid according to Claim 1, which comprises the following steps:

(a) carrying out an entirely chemical synthesis in a manner known per se, or

(b) chemically synthesizing oligonucleotides, labelling the oligonucleotides, hybridizing the oligonucleotides to the DNA of an insect cDNA library, selecting positive clones and isolating the hybridizing DNA from positive clones, or

(c) chemically synthesizing oligonucleotides and amplifying the target DNA by means of PCR.

17. Regulatory region which naturally controls transcription of a nucleic acid according to Claim 1 in insect cells and ensures specific expression.

18. Process for discovering novel active compounds for plant protection, in particular compounds which alter the conducting properties of receptors according to Claim 9, which comprises the following steps:

(a) providing a host cell according to one of Claims 4 to 7,

(b) culturing the host cell in the presence of a compound or a sample which comprises a multiplicity of compounds, and

(c) detecting altered receptor properties.

19. Process for discovering a compound which binds to receptors according to Claim 9, which encompasses the following steps:

(a) bringing a host cell according to one of Claims 4 to 7, a polypeptide according to Claim 8 or a receptor according to Claim 9 into contact with a compound or a mixture of compounds under conditions which permit interaction of the compound(s) with the host cell, the polypeptide or the receptor, and

(b) determining the compound(s) which bind(s) specifically to the receptors.

20. Process for discovering compounds which alter the expression of receptors according to Claim 9, which comprises the following steps:

(a) bringing a host cell according to one of Claims 4 to 7 or a transgenic invertebrate according to Claim 11 or 12 into contact with a compound or a mixture of compounds,

(b) determining the receptor concentration, and

(c) determining the compound(s) which specifically influence(s) the expression of the receptor.

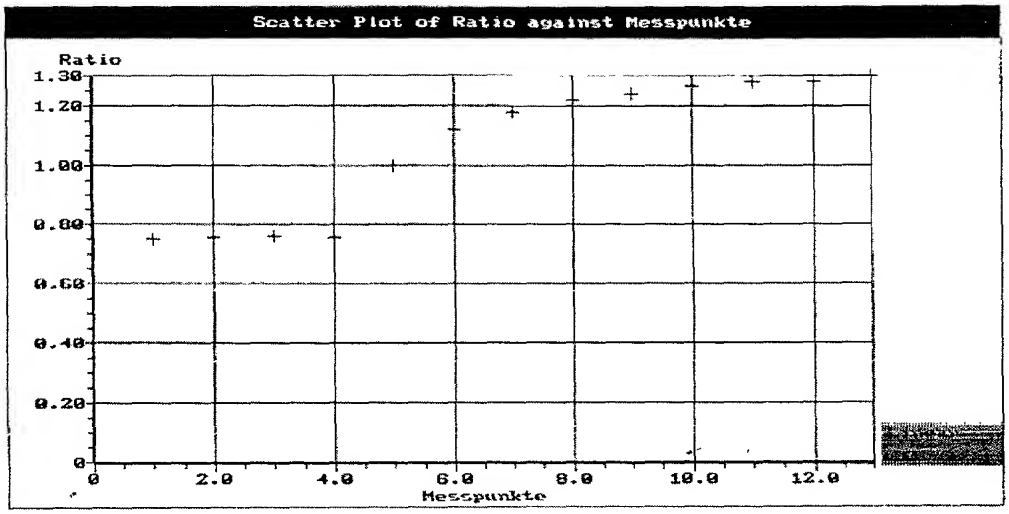
21. Use of at least one nucleic acid according to Claim 1, one vector according to Claim 2 or 3, one regulatory region according to Claim 16 or one antibody according to Claim 11 for discovering novel active compounds for plant protection or for discovering genes which encode polypeptides which are involved in synthesizing functionally similar acetylcholine receptors in insects.

Nucleic acids which encode insect acetylcholine receptor subunits

Abstract

The invention relates to nucleic acids which encode insect acetylcholine receptor subunits, to the corresponding polypeptides, and to processes for discovering novel active compounds for plant protection.

Figure 1



Measuring Points

SEQUENCE LISTING

<110> Bayer Aktiengesellschaft

<120> Nucleic acids which encode
insect acetylcholine receptor subunits

<130> Le A 33 020-Foreign Countries

<140>

<141>

<150> DE-198 19 829.9

<151> 1998-05-04

<160> 6

<170> PatentIn Ver. 2.1

<210> 1

<211> 2886

<212> DNA

<213> Drosophila melanogaster

<220>

<221> CDS

<222> (372)..(2681)

<400> 1

ggcacgagaa aaagttgtgg tataaacttt tattgtagga aaacgcataa aaataataga 60

aaaacgctct tcgggttgta aagaaaataa gaagacaaaa gaaagacatg aaaacgttgc 120

aaacaataaa gcatatactt gccatattga tataaaggga aatcgtgaaa aggcggtgaa 180

aatttcgtaa gattagttgg tattaagggc agcccatgca cacagctaaa aagggaacta 240

aaaaaacccc gcacagaaca atgaaagctg cagcagctgg ataaggccga caaaaccgaa 300

aattatatta ttgtaatcta gtagagagca gacaacatat ccgctggcaa caaccaacac 360

cgaaagagac t atg aaa aat gca caa ctg aaa ctg act gaa gtt gac gat 410
Met Lys Asn Ala Gln Leu Lys Leu Thr Glu Val Asp Asp
1 5 10

gat gag ctg tgg ctg gca gta aga tta gcg cac tgc agc agc aac ttt 458
Asp Glu Leu Trp Leu Ala Val Arg Leu Ala His Cys Ser Ser Asn Phe
15 20 25

agc agc agt agc agc aca aga acc acc agc agc aac cag agg cac aac 506
Ser Ser Ser Ser Ser Thr Arg Thr Thr Ser Ser Asn Gln Arg His Asn
30 35 40 45

cag caa ctc aca aca ctg caa cca agg agc tta agt aca aaa cac cac 554
Gln Gln Leu Thr Thr Leu Gln Pro Arg Ser Leu Ser Thr Lys His His
50 55 60

agc aac att gca agc gag cag cac aat agc cag caa cag gag cca gca 602
Ser Asn Ile Ala Ser Glu Gln His Asn Ser Gln Gln Gln Glu Pro Ala
65 70 75

tcg aag gac gag gat gta gcc aac cac ggt aga agc aat gac cag cag 650
Ser Lys Asp Glu Asp Val Ala Asn His Gly Arg Ser Asn Asp Gln Gln

80 85 90

acg cat ctg caa cag cta gac agc agc aac atg ttg tcg cca aag aca 698
Thr His Leu Gln Gln Leu Asp Ser Ser Asn Met Leu Ser Pro Lys Thr
95 100 105

gcc gca gca gca act gct gcc ggc gat gaa gca aca acc caa caa cca 746
Ala Ala Ala Ala Thr Ala Ala Gly Asp Glu Ala Thr Thr Gln Gln Pro
110 115 120 125

aca aac ata aga ctg tgt gca cgc aag cga caa cga ttg cgt cgc cga 794
Thr Asn Ile Arg Leu Cys Ala Arg Lys Arg Gln Arg Leu Arg Arg Arg
130 135 140

cga aaa aga aaa cca gca acc cca aac gaa aca gat atc aag aaa caa 842
Arg Lys Arg Lys Pro Ala Thr Pro Asn Glu Thr Asp Ile Lys Lys Gln
145 150 155

cag caa ctt agc atg cct ccc ttc aaa acg cgc aaa tcc acg gac acc 890
Gln Gln Leu Ser Met Pro Pro Phe Lys Thr Arg Lys Ser Thr Asp Thr
160 165 170

tac agc aca cca gca gca aca acc agc tgt ccg aca gcc acc tac atg 938
Tyr Ser Thr Pro Ala Ala Thr Thr Ser Cys Pro Thr Ala Thr Tyr Met
175 180 185

caa tgt cga gcc agc gac aat gag ttc agt att ccg ata tcg aga cat 986
Gln Cys Arg Ala Ser Asp Asn Glu Phe Ser Ile Pro Ile Ser Arg His
190 195 200 205

gat aga gta tcc acg gcc aca ttc gcc tgg gtg ttg cat gtg ctg cag 1034

```
Asp Arg Val Ser Thr Ala Thr Phe Ala Trp Val Leu His Val Leu Gln
                210                 215                 220

gtg ctg ctc gtg tcg ctg caa cag tgg caa ctt cac gtg caa cag cga   1082
Val Leu Leu Val Ser Leu Gln Gln Trp Gln Leu His Val Gln Gln Arg
            225                 230                 235

tcg gtg cta ctg ttc aga agg atc gca gcg agc acc atc gcc ttc att   1130
Ser Val Leu Leu Phe Arg Arg Ile Ala Ala Ser Thr Ile Ala Phe Ile
        240                 245                 250

tcc tat tta ggc agc ttt gca gcg caa ctg aaa aat agc agc agc agc   1178
Ser Tyr Leu Gly Ser Phe Ala Ala Gln Leu Lys Asn Ser Ser Ser Ser
    255                 260                 265

agt agc agc agc aac agc agc aac aac agc agc acg caa ata tta aac   1226
Ser Ser Ser Ser Asn Ser Ser Asn Asn Ser Ser Thr Gln Ile Leu Asn
270                 275                 280                 285

gga ctt aat aaa cac tca tgg ata ttt tta ttg ata tat ttg aat tta   1274
Gly Leu Asn Lys His Ser Trp Ile Phe Leu Leu Ile Tyr Leu Asn Leu
                290                 295                 300

tct gct aaa gtt tgc cta gca gga tat cat gaa aag aga ctg tta cac   1322
Ser Ala Lys Val Cys Leu Ala Gly Tyr His Glu Lys Arg Leu Leu His
            305                 310                 315

gat ctt ttg gat cct tat aat aca cta gaa cgt ccc gtt ctc aat gaa   1370
Asp Leu Leu Asp Pro Tyr Asn Thr Leu Glu Arg Pro Val Leu Asn Glu
        320                 325                 330
```

tcg gac ccg tta caa tta agc ttt ggt tta act tta atg caa att atc 1418
Ser Asp Pro Leu Gln Leu Ser Phe Gly Leu Thr Leu Met Gln Ile Ile
335 340 345

gat gtg gac gag aaa aat caa ttg cta gtc act aat gtg tgg tta aaa 1466
Asp Val Asp Glu Lys Asn Gln Leu Leu Val Thr Asn Val Trp Leu Lys
350 355 360 365

ctg gag tgg aac gac atg aat ctc cgc tgg aac acc tcc gac tat ggc 1514
Leu Glu Trp Asn Asp Met Asn Leu Arg Trp Asn Thr Ser Asp Tyr Gly
370 375 380

gga gtt aag gat ctg cga ata ccg ccg cat cgc atc tgg aag ccg gac 1562
Gly Val Lys Asp Leu Arg Ile Pro Pro His Arg Ile Trp Lys Pro Asp
385 390 395

gtg ctg atg tac aac agt gcg gat gag gga ttt gac ggc acc tac cag 1610
Val Leu Met Tyr Asn Ser Ala Asp Glu Gly Phe Asp Gly Thr Tyr Gln
400 405 410

acg aac gtg gtg gtg cgg aac aac ggc tcg tgt cta tac gtt ccg ccg 1658
Thr Asn Val Val Val Arg Asn Asn Gly Ser Cys Leu Tyr Val Pro Pro
415 420 425

ggg atc ttc aag tcg acg tgc aag atc gac atc acg tgg ttc ccc ttc 1706
Gly Ile Phe Lys Ser Thr Cys Lys Ile Asp Ile Thr Trp Phe Pro Phe
430 435 440 445

gat gac cag cgg tgc gag atg aag ttc ggc agt tgg acc tac gac gga 1754
Asp Asp Gln Arg Cys Glu Met Lys Phe Gly Ser Trp Thr Tyr Asp Gly
450 455 460

ttc cag ctg gat tta caa tta caa gat gaa act ggc ggt gat atc agc 1802
Phe Gln Leu Asp Leu Gln Leu Gln Asp Glu Thr Gly Gly Asp Ile Ser
465 470 475

agt tac gtg ctc aac ggc gag tgg gaa cta ctg ggt gtg ccc ggc aaa 1850
Ser Tyr Val Leu Asn Gly Glu Trp Glu Leu Leu Gly Val Pro Gly Lys
480 485 490

cgt aac gag atc tat tac aac tgc tgc ccg gaa ccc tat ata gac atc 1898
Arg Asn Glu Ile Tyr Tyr Asn Cys Cys Pro Glu Pro Tyr Ile Asp Ile
495 500 505

acc ttc gcc atc atc atc cgc cga cga aca ctg tac tat ttc ttc aac 1946
Thr Phe Ala Ile Ile Ile Arg Arg Arg Thr Leu Tyr Tyr Phe Phe Asn
510 515 520 525

ctg atc ata cct tgt gta ctg att gcc tcc atg gcc ttg ctc gga ttc 1994
Leu Ile Ile Pro Cys Val Leu Ile Ala Ser Met Ala Leu Leu Gly Phe
530 535 540

acc ctg ccg cca gat tcg ggt gaa aaa tta tcg ctg ggt gtt acc atc 2042
Thr Leu Pro Pro Asp Ser Gly Glu Lys Leu Ser Leu Gly Val Thr Ile
545 550 555

ttg ctc tcg ctg acc gtg ttt ctg aat atg gtt gcc gag aca atg ccg 2090
Leu Leu Ser Leu Thr Val Phe Leu Asn Met Val Ala Glu Thr Met Pro
560 565 570

gct act tcc gat gcg gtg cca ttg tgg ata cgc atc gtg ttt ttg tgc 2138
Ala Thr Ser Asp Ala Val Pro Leu Trp Ile Arg Ile Val Phe Leu Cys

575 580 585

tgg ctg cca tgg ata ttg cga atg agt cgc cca gga cga ccg ctg atc 2186
Trp Leu Pro Trp Ile Leu Arg Met Ser Arg Pro Gly Arg Pro Leu Ile
590 595 600 605

cta gag ttc ccg acc acg ccc tgt tcg gac aca tcc tcc gag cgg aag 2234
Leu Glu Phe Pro Thr Thr Pro Cys Ser Asp Thr Ser Ser Glu Arg Lys
610 615 620

cac cag ata ctc tcc gac gtt gag ctg aaa gag cgc tcg tcg aaa tcg 2282
His Gln Ile Leu Ser Asp Val Glu Leu Lys Glu Arg Ser Ser Lys Ser
625 630 635

ctg ctg gcc aac gta cta gac atc gat gat gac ttc cgg cac aat tgt 2330
Leu Leu Ala Asn Val Leu Asp Ile Asp Asp Asp Phe Arg His Asn Cys
640 645 650

cgc ccc atg acg ccc ggc gga aca ctg cca cac aac ccg gct ttc tat 2378
Arg Pro Met Thr Pro Gly Gly Thr Leu Pro His Asn Pro Ala Phe Tyr
655 660 665

cgc acg gtt tat gga caa ggc gac gat ggc agc att ggg cca att ggc 2426
Arg Thr Val Tyr Gly Gln Gly Asp Asp Gly Ser Ile Gly Pro Ile Gly
670 675 680 685

agc acc cga atg ccg gat gcg gtc acc cat cat acg tgc atc aaa tca 2474
Ser Thr Arg Met Pro Asp Ala Val Thr His His Thr Cys Ile Lys Ser
690 695 700

tca act gaa tat gaa tta ggt tta atc tta aag gaa att cgc ttt ata 2522

Ser Thr Glu Tyr Glu Leu Gly Leu Ile Leu Lys Glu Ile Arg Phe Ile
705 710 715

act gat cag cta cgt aaa gat gac gag tgc aat gac att gcc aat gat 2570
Thr Asp Gln Leu Arg Lys Asp Asp Glu Cys Asn Asp Ile Ala Asn Asp
720 725 730

tgg aaa ttt gca gct atg gtc gtt gac aga ctg tgc ctt atc ata ttc 2618
Trp Lys Phe Ala Ala Met Val Val Asp Arg Leu Cys Leu Ile Ile Phe
735 740 745

aca atg ttc gca ata tta gcc aca ata gct gta cta cta tcg gca cca 2666
Thr Met Phe Ala Ile Leu Ala Thr Ile Ala Val Leu Leu Ser Ala Pro
750 755 760 765

cat att att gtc tcg tagccatatg ggcgaggtgg ttattgttat tggttttatt 2721
His Ile Ile Val Ser
770

ataaaatcaa tttgttaatt attaaattaa taacgaaact ctttaagtaa attaaaacta 2781

aaaagacact aaaaaagcac aaaaaaatag gaaaatacat gataaaaccc atgaactaaa 2841

taatacatcc aagaaaaacc aaaacaaaaa aaaaaaaaaa aaaaa 2886

<210> 3

<211> 3701

<212> DNA

<213> Heliothis virescens

<220>

<221> CDS

<222> (335)..(1822)

<400> 3

ggcacgagcc gctgccccac ggtcggccgc actccgctga acaacaatgc tcaaaaacac 60

gccgtgactc cacacacatc ccctcggcgc agtaggcgat gtttgaggat cggacggcac 120

gcgtggccgt cggcgagcgg tcgtgaacaa gttgcataca tatgaaaacc gtaaaaagat 180

tgaattttaa gccgatcgtg ttcgatagat cctaatagag aagcgggagt gcggcgtttg 240

gtaggcgggg gtcgagtcgc gcggtcgggg gaaatggcgc ggcgcggggc ggcggcggcg 300

gcggcgcgcg gcgcggcggc gtcgcggcgc tgac atg ggc ggg cgg gcg cgc cgc 355
Met Gly Gly Arg Ala Arg Arg
1 5

tcg cac ttg gcg gcg ccc gcg ggc ctg ctg ctg ctg ctg tgc ctg ctc 403
Ser His Leu Ala Ala Pro Ala Gly Leu Leu Leu Leu Cys Leu Leu
10 15 20

tgg ccg agg ggg gca cgc tgc ggg tac cac gag aag cgg cta ctg cac 451
Trp Pro Arg Gly Ala Arg Cys Gly Tyr His Glu Lys Arg Leu Leu His
25 30 35

```
cac cta ttg gac cac tac aac gta ctg gag agg ccc gtc gtc aac gag   499
His Leu Leu Asp His Tyr Asn Val Leu Glu Arg Pro Val Val Asn Glu
 40                  45                  50                  55

agc gac ccg ctg cag ctc tcc ttc ggc ctc acg ctc atg cag atc atc   547
Ser Asp Pro Leu Gln Leu Ser Phe Gly Leu Thr Leu Met Gln Ile Ile
                 60                  65                  70

gac gtg gac gag aag aac cag ctt tta ata aca aac atc tgg cta aaa   595
Asp Val Asp Glu Lys Asn Gln Leu Leu Ile Thr Asn Ile Trp Leu Lys
             75                  80                  85

cta gag tgg aat gat atg aac ttg agg tgg aac act tca gat ttc ggc   643
Leu Glu Trp Asn Asp Met Asn Leu Arg Trp Asn Thr Ser Asp Phe Gly
         90                  95                 100

ggg gtc aaa gat tta aga gtg cca ccc cac aga cta tgg aaa cca gac   691
Gly Val Lys Asp Leu Arg Val Pro Pro His Arg Leu Trp Lys Pro Asp
     105                 110                 115

gtc ctt atg tac aac agc gcg gac gaa ggg ttc gac agc acg tat cca   739
Val Leu Met Tyr Asn Ser Ala Asp Glu Gly Phe Asp Ser Thr Tyr Pro
120                 125                 130                 135

acg aac gtg gtg gtg cgg aac aac ggc tcg tgt ctg tac gtg ccg ccc   787
Thr Asn Val Val Val Arg Asn Asn Gly Ser Cys Leu Tyr Val Pro Pro
                140                 145                 150

ggc atc ttc aag agc acc tgc aag atc gac atc acc tgg ttc ccc ttc   835
Gly Ile Phe Lys Ser Thr Cys Lys Ile Asp Ile Thr Trp Phe Pro Phe
```

155 160 165

gac gac caa cga tgc gag atg aag ttt ggc agc tgg act tat gat ggt 883
Asp Asp Gln Arg Cys Glu Met Lys Phe Gly Ser Trp Thr Tyr Asp Gly
170 175 180

tat cag ttg gat cta caa cta cag gat gaa ggg ggc gga gat ata agc 931
Tyr Gln Leu Asp Leu Gln Leu Gln Asp Glu Gly Gly Gly Asp Ile Ser
185 190 195

agt ttt gtc acg aat ggc gaa tgg gag tta ata gga gtc ccc ggc aag 979
Ser Phe Val Thr Asn Gly Glu Trp Glu Leu Ile Gly Val Pro Gly Lys
200 205 210 215

cgc aac gag atc tac tac aac tgt tgt ccg gag cca tac atc gac atc 1027
Arg Asn Glu Ile Tyr Tyr Asn Cys Cys Pro Glu Pro Tyr Ile Asp Ile
220 225 230

acg ttt gcg gtg gtg atc cgg agg aaa acg ctc tac tac ttc ttc aat 1075
Thr Phe Ala Val Val Ile Arg Arg Lys Thr Leu Tyr Tyr Phe Phe Asn
235 240 245

ctg atc gtg ccc tgc gtg ctc atc gcc tcc atg gct cta ttg ggg ttc 1123
Leu Ile Val Pro Cys Val Leu Ile Ala Ser Met Ala Leu Leu Gly Phe
250 255 260

acc ttg cct cca gac tcc gga gaa aag ttg tct tta ggt gtg acg ata 1171
Thr Leu Pro Pro Asp Ser Gly Glu Lys Leu Ser Leu Gly Val Thr Ile
265 270 275

tta ctg tcg ttg acg gtg ttc ctc aac atg gtg gcg gag acg atg cca 1219

```
Leu Leu Ser Leu Thr Val Phe Leu Asn Met Val Ala Glu Thr Met Pro
280                 285                 290                 295

gcg acg tcg gac gcc gtg ccc ttg ctc ggc acc tac ttc aac tgc atc   1267
Ala Thr Ser Asp Ala Val Pro Leu Leu Gly Thr Tyr Phe Asn Cys Ile
                300                 305                 310

atg ttc atg gtg gct tcc tcc gtc gtc tcc acc ata ctg atc ctc aac   1315
Met Phe Met Val Ala Ser Ser Val Val Ser Thr Ile Leu Ile Leu Asn
            315                 320                 325

tac cac cac cgg cac gca gac act cac gaa atg agt gat tgg att cgt   1363
Tyr His His Arg His Ala Asp Thr His Glu Met Ser Asp Trp Ile Arg
        330                 335                 340

tgc gtg ttc ctt tat tgg ctg ccg tgg gtg ctg cgc atg tca cgg ccc   1411
Cys Val Phe Leu Tyr Trp Leu Pro Trp Val Leu Arg Met Ser Arg Pro
    345                 350                 355

ggc tcg gcg acg acg ccg ccg ccg gcg cgc gta cct ccg ccg ccg gac   1459
Gly Ser Ala Thr Thr Pro Pro Pro Ala Arg Val Pro Pro Pro Pro Asp
360                 365                 370                 375

ctg gag ctg cgc gag cgc tcc tcc aag tcg ctc cta gcg aac gtg ctc   1507
Leu Glu Leu Arg Glu Arg Ser Ser Lys Ser Leu Leu Ala Asn Val Leu
                380                 385                 390

gac atc gat gac gac ttc cgc cac ccg caa gcg cag cag ccg caa tgc   1555
Asp Ile Asp Asp Asp Phe Arg His Pro Gln Ala Gln Gln Pro Gln Cys
            395                 400                 405
```

tgc cga tac tac agg ggg ggt gag gag aat ggc gcg ggg ttg gcg gcg 1603
Cys Arg Tyr Tyr Arg Gly Gly Glu Glu Asn Gly Ala Gly Leu Ala Ala
410 415 420

cac agt tgc ttc ggt gtc gac tac gag ctc tcc ctc att ctg aag gag 1651
His Ser Cys Phe Gly Val Asp Tyr Glu Leu Ser Leu Ile Leu Lys Glu
425 430 435

att aga gtc atc aca gat cag atg cgc aag gac gac gaa gat gcg gac 1699
Ile Arg Val Ile Thr Asp Gln Met Arg Lys Asp Asp Glu Asp Ala Asp
440 445 450 455

att tcg cgc gac tgg aag ttc gcc gcc atg gtc gtg gac aga ctg tgc 1747
Ile Ser Arg Asp Trp Lys Phe Ala Ala Met Val Val Asp Arg Leu Cys
460 465 470

ctt att atc ttt acc ctg ttc aca atc atc gcc acg cta gcc gtg ctg 1795
Leu Ile Ile Phe Thr Leu Phe Thr Ile Ile Ala Thr Leu Ala Val Leu
475 480 485

ctg tcc gcg cca cac atc atg gtg tcg tagcgacccg cccgcttgcg 1842
Leu Ser Ala Pro His Ile Met Val Ser
490 495

gatacgcatg cgaaaagttc tgtgataccg cgaatatttg ttaagttgtg atgagcgaag 1902

tggcgcggac ggtgacgccg cggcgtcgga gttgccgccg cctgcctcgc cgcccgcgcc 1962

cccctgtaga cataagttac cgctgactgc caaccctgta cgttcaacaa ataactgccc 2022

atccgactaa cgtcttttat ccccttgaaa aattcagcga ttgtgtaccc ctttcttcca 2082

agaatacaat gacaaatggt cgtcacgctc agtggaatca atcccgtact cttcgcccga 2142

tatttccctt agggtatgtc acgagtttga atgagcggtt ccgtatcaga cgttccgtcc 2202

ccggaacggt cgtcccctgc gataaagtgg cagtacgtgc tatacaggca cttaaggccg 2262

ccacgccacg gcgccgcggt gcgctcgggc cgcgaacccg cgaccctcac cgctgcaagt 2322

ggccacccac tagacaagac tgcggcagaa aatatttgca caaaaacgtc ttccttctta 2382

ccgatgaacg acctgattcg catttaaaat taaactttgt tagaacttct tcgattcttg 2442

aaatctattg tacagtttag agtttgggcg gtgaaacaat ggccctttgt ttccttcttg 2502

ttcgattcca tgaatcgtgg ttataatccc tagttttatt ttcggatata tttgtgtcag 2562

tagctagtat agaactttac aaacaatgtt gattcaattg gtacaggttg tgatatgcct 2622

cgttgtgaac gggtccgata ttgttataaa tggtaaaata cccatggcta tagcttaata 2682

aatcgttcgt taaaagttgt agttaaacaa atattatttt aataaagtca tatctgggtc 2742

ttccggaacg acttttacaa ataattaaat tacatattaa tatcacgttt gtacttcttt 2802

ccatacagtt acagtaattc gtatgctgaa aataatatta gcttgtaaaa ttttcttctt 2862

cgaaaattta ttcaaacaga tgcgaccatc gtttcaaaca tttacatgta atatagaact 2922

cattttataa gatatacaac attttataag tacaagaagt tgtaacatga accggttttt 2982

cgttacatag agggtataac acaaaggtgc ctacatattg acagatgcga agcacgatca 3042

gttgataagc acaggtacac tatatcctga catccgacag tcctgccgct cgtctgccac 3102

actcggaaac attcgacagt tcagtttact gctccgccat catcgattgt taagtttgtt 3162

gttctaactc atcgcattca tttcattcaa aaacattgta aacctctcaa ggggaaaacg 3222

tgttgtaaac agtgagagtg cgcgggtaca accgacacgc gaatgtaccc tcgcaaggct 3282

cctgtaatgt tttcctcttc cgaggtgttg ctgagagtaa tcttagacgg tccgatggaa 3342

gttgcggacc ggatatgatt acaagtcaat gtttttaagt catccgttta tttattgtta 3402

tatcttctta ccattcgcta gaggttgtgt gacgacccgg acggtgggcg ccgcaacccg 3462

cacacgcggg gttccatctt tgtattagat ggaagttgtg cggcatctct ccgtcggcaa 3522

tgggacaacc cgttgtcccc aacatttgtt caattgttag ggttaactct gaattgcact 3582

ttgtttatta aatataaacg aatgaaacaa aaaaaaaaaa aaaaaactcg agagtacttc 3642

tagagcggcc gcgggcccat cgattttcca cccgggtggg gtaccargta agtgtaccc 3701

<210> 5

<211> 3109

<212> DNA

<213> Heliothis virescens

<220>

<221> CDS

<222> (95)..(1597)

<400> 5

ggcacgagcc ggccgcacgt tgtcccaggc cgcatgagcg cgccggcgtg ctagcgcagc 60

gtgcgcgggt gtggtatgcc cgcgcgtcgc cgct atg gcc cct atg ttg gcg gcc 115
Met Ala Pro Met Leu Ala Ala
1 5

ttg gcg ctg ctg gct ttg ctg ccc gta tcg gag caa ggt cct cac gag 163
Leu Ala Leu Leu Ala Leu Leu Pro Val Ser Glu Gln Gly Pro His Glu
10 15 20

aag aga ctc ctg aac gcg ttg ctg gcg aac tac aac acc ctg gag cga 211
Lys Arg Leu Leu Asn Ala Leu Leu Ala Asn Tyr Asn Thr Leu Glu Arg
25 30 35

ccg gtg gcc aac gag agc gaa ccg cta gag gtc agg ttc ggc ttg acc 259
Pro Val Ala Asn Glu Ser Glu Pro Leu Glu Val Arg Phe Gly Leu Thr
40 45 50 55

ttg cag caa atc att gac gtg gac gag aag aat caa cta ctt ata acc 307

Leu Gln Gln Ile Ile Asp Val Asp Glu Lys Asn Gln Leu Leu Ile Thr
60 65 70

aat ata tgg ctg tcg ttg gag tgg aat gac tac aac ctg agg tgg aac 355
Asn Ile Trp Leu Ser Leu Glu Trp Asn Asp Tyr Asn Leu Arg Trp Asn
75 80 85

gac agc gag tat ggc ggg gtc aag gac ctc agg atc acg ccc aac aag 403
Asp Ser Glu Tyr Gly Gly Val Lys Asp Leu Arg Ile Thr Pro Asn Lys
90 95 100

ttg tgg aag ccg gac gtc ctt atg tat aat agt gct gac gag ggt ttt 451
Leu Trp Lys Pro Asp Val Leu Met Tyr Asn Ser Ala Asp Glu Gly Phe
105 110 115

gac ggg acc tac cag acc aac gtg gtg gtc aga agc ggc ggc agt tgc 499
Asp Gly Thr Tyr Gln Thr Asn Val Val Val Arg Ser Gly Gly Ser Cys
120 125 130 135

ctg tac gtg cca cct ggc ata ttc aag agc aca tgc aag atg gac atc 547
Leu Tyr Val Pro Pro Gly Ile Phe Lys Ser Thr Cys Lys Met Asp Ile
140 145 150

gcg tgg ttt ccc ttc gac gac caa cac tgt gat atg aag ttc ggt agc 595
Ala Trp Phe Pro Phe Asp Asp Gln His Cys Asp Met Lys Phe Gly Ser
155 160 165

tgg aca tat gac ggc aat cag ttg gat ctg gtg cta aaa gat gag gca 643
Trp Thr Tyr Asp Gly Asn Gln Leu Asp Leu Val Leu Lys Asp Glu Ala
170 175 180

```
ggc ggc gat cta tcg gac ttc ata aca aat ggg gag tgg tat cta ata   691
Gly Gly Asp Leu Ser Asp Phe Ile Thr Asn Gly Glu Trp Tyr Leu Ile
    185                 190                 195

gga atg cca ggc aaa aag aac aca ata aca tac gcg tgc tgc ccc gag   739
Gly Met Pro Gly Lys Lys Asn Thr Ile Thr Tyr Ala Cys Cys Pro Glu
200                 205                 210                 215

ccc tac gtg gac gtc acc ttc acc atc atg ata aga aga cga acc ttg   787
Pro Tyr Val Asp Val Thr Phe Thr Ile Met Ile Arg Arg Arg Thr Leu
                220                 225                 230

tac tac ttc ttc aac ctg atc gtc ccg tgc gtg ctg atc tca tcg atg   835
Tyr Tyr Phe Phe Asn Leu Ile Val Pro Cys Val Leu Ile Ser Ser Met
            235                 240                 245

gca ctc ctc ggc ttc aca ctg cca cca gac tcc gga gag aaa ctc aca   883
Ala Leu Leu Gly Phe Thr Leu Pro Pro Asp Ser Gly Glu Lys Leu Thr
        250                 255                 260

ctt gga gtc act att ctt cta tcg ctg acg gtg ttc ctc aac ctg gta   931
Leu Gly Val Thr Ile Leu Leu Ser Leu Thr Val Phe Leu Asn Leu Val
    265                 270                 275

gcc gag acc ctg cca cag gtc tcc gac gct atc ccc ctg tta ggg acg   979
Ala Glu Thr Leu Pro Gln Val Ser Asp Ala Ile Pro Leu Leu Gly Thr
280                 285                 290                 295

tac ttc aat tgc atc atg ttc atg gta gcg tcg tct gtg gta ctg act   1027
Tyr Phe Asn Cys Ile Met Phe Met Val Ala Ser Ser Val Val Leu Thr
                300                 305                 310
```

```
gtg gtg gta ctc aat tac cac cat cga aca gct gat ata cat gaa atg   1075
Val Val Val Leu Asn Tyr His His Arg Thr Ala Asp Ile His Glu Met
            315                 320                 325

cca cag tgg ata aaa tca gta ttc cta caa tgg ttg cca tgg ata ctg   1123
Pro Gln Trp Ile Lys Ser Val Phe Leu Gln Trp Leu Pro Trp Ile Leu
        330                 335                 340

cga atg tcg agg cca ggg aag aag atc acc agg aag act ata atg atg   1171
Arg Met Ser Arg Pro Gly Lys Lys Ile Thr Arg Lys Thr Ile Met Met
    345                 350                 355

aac acg agg atg agg gag ctg gaa ctg aag gag agg tcg tcg aag tcc   1219
Asn Thr Arg Met Arg Glu Leu Glu Leu Lys Glu Arg Ser Ser Lys Ser
360                 365                 370                 375

ttg ctg gcg aat gtt cta gat att gat gat gac ttc aga cac ggc cct   1267
Leu Leu Ala Asn Val Leu Asp Ile Asp Asp Asp Phe Arg His Gly Pro
                380                 385                 390

ccg cct cct aac agt act gcc tcg acc ggg aat ttg gga cct ggg tgc   1315
Pro Pro Pro Asn Ser Thr Ala Ser Thr Gly Asn Leu Gly Pro Gly Cys
            395                 400                 405

tca ata ttc cgc acg gat ttc cgt cgg tcg ttc gtc cgt ccg tcc acg   1363
Ser Ile Phe Arg Thr Asp Phe Arg Arg Ser Phe Val Arg Pro Ser Thr
        410                 415                 420

atg gaa gac gtg ggc ggc ggg ctg ggt agc cac cat cgc gag ctg cac   1411
Met Glu Asp Val Gly Gly Gly Leu Gly Ser His His Arg Glu Leu His
```

425 430 435

ctc ata ctg aga gag ctg cag ttc atc acg gcc agg atg aag aag gct 1459
Leu Ile Leu Arg Glu Leu Gln Phe Ile Thr Ala Arg Met Lys Lys Ala
440 445 450 455

gat gag gaa gcc gag ctg atc agc gac tgg aag ttt gct gcg atg gtt 1507
Asp Glu Glu Ala Glu Leu Ile Ser Asp Trp Lys Phe Ala Ala Met Val
460 465 470

gtt gat agg ttt tgc ctg ttc gtg ttc aca ctt ttc aca atc atc gcg 1555
Val Asp Arg Phe Cys Leu Phe Val Phe Thr Leu Phe Thr Ile Ile Ala
475 480 485

aca gta gct gtc ctg tta tcg gca ccg cat atc atc gtg caa 1597
Thr Val Ala Val Leu Leu Ser Ala Pro His Ile Ile Val Gln
490 495 500

tgaaccaacc actgagccgg caactccggc gcatgaatga gagaaataat tattagatcg 1657

ccgatttgta attataattg ataatgtaat taaattaaat acgtggttga aacgcacacg 1717

tctccataac aaagtcttaa gacattaaat tatgataaat ttacatattg tagttaagtc 1777

gagtgttgat ggaaatttta gccggcgcaa ggagtttcgt gaaggtctgt atatattttt 1837

tcttattgtt gtatattgta tcgttgttca tgttttcttt caggaagtga gctttgtact 1897

gtttgtttct tcgatggcag gtgcacttca gttcaggctg aaatttccat taacatttat 1957

ttaaacaaat gtgatgttga ctaggatgtt atacagataa atgttgacgt gtataatttg 2017

ttaaaataaa caatattaat tactattact aaacgatatt ataaacgaag tactaacgag 2077

ggttacttta atgggaagaa cgctaagctg gcacagagtt gcattaattt gaaaaaagaa 2137

attacggaaa aaagtttatt gaaaattgaa ctttttggaa ggaaagtaac gtttgatcaa 2197

aaaagtttgt aaaacgaaag ttcggttctg cgccaatact ggaattaaaa ttctcgtaaa 2257

tattagggaa aagaaggtcc tttaaaacaa aagatttgaa ccggcatcct ttttacaagt 2317

aatgagggat cacagatgat gacaaaaaac cttagggtat ataagtaatg tacataatgg 2377

atcaaatatc ggtagagtca agaatagtta acgatttaag attattccat tcgatattaa 2437

aattcgatta gcgattgtcg ctgcgtctac tttgatacat atcgatttga atcgatattg 2497

tataaattta gatagatcgg acattagtaa tgagtatgga cgttttaatt tttaaaaaag 2557

aatgtactac gaagattaaa tccaggaatt gttaaacagt tatggaattg ataagaaatc 2617

aacaattaat acggaaccaa aggtagacta ggtgtagcat caggagattg aattaaaaca 2677

taaattagga ccgacttaaa tggaacttgc gagtgtattg ataacttttt aatttaaaaa 2737

ctcattgtcg attaaatgga gaataacttt tgatctctcg tatcgataaa tgctcactta 2797

actatcgata gcgtaatatt ataactgtta gtatatcgat atgggagtaa gtcactagca 2857

tcagaaatag tcattaatta ggaatcggtt tgtgttaatg ttatgcttag cgaaaatatt 2917

acaatgctgt tgatatcact aaccatcacg taaccatatt gataaaatgt aaatacagaa 2977

tattgcggtg tgtatttgta tataaatttt agaaaaaaaa aaaaaaaaaa aactcgagag 3037

tacttctaga gcggccgcgg gcccatcgat tttccacccg ggtggggtac caggtaagtg 3097

tacccaattc gc 3109